Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

Benchmarking Operational Risk Stress Testing Models

Filippo Curti, Marco Migueis, and Robert Stewart

2019-038

Please cite this paper as:
Curti, Filippo, Marco Migueis, and Robert Stewart (2019). "Benchmarking Operational Risk Stress Testing Models," Finance and Economics Discussion Series 2019-038. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2019.038.

Benchmarking Operational Risk Stress Testing Models[1]
September 6, 2018

Filippo Curti[2]
Federal Reserve Bank of Richmond

Marco Migueis[3]
Federal Reserve Board

Robert Stewart[4]

Abstract

The Federal Reserve's Comprehensive Capital Analysis and Review (CCAR) requires large bank holding companies (BHCs) to project losses under stress scenarios. In this paper, we propose multiple benchmarks for operational loss projections and document the industry distribution relative to these benchmarks. The proposed benchmarks link BHCs' loss projections with both financial characteristics and metrics of historical loss experience. These benchmarks capture different measures of exposure and together provide a comprehensive view of the reasonability of model outcomes. Furthermore, we employ several approaches to assess the conservatism of BHCs' stress loss projections and our estimates for the conservatism of loss projections for the median bank range from the 90th percentile to above the 99th percentile of the operational loss distribution.

JEL Classification: G21, G28, G32

Keywords: Operational Risk, Stress Testing, Benchmarking

[1] The views expressed in this manuscript belong to the authors and do not represent official positions of the Federal Reserve Board, the Federal Reserve Bank of Richmond, or the Federal Reserve System.
[2] Email: filippo.curti@rich.frb.org.
[3] Email: marco.a.migueis@frb.gov.
[4] Email: roberttstewart.15@gmail.com.

I – Introduction

Benchmarking is a model validation technique whereby model outputs are compared to alternative models or metrics. Federal Reserve guidance recommends that banks use benchmarking to assess the efficacy of their models (Board of Governors of the Federal Reserve System 2011). One important purpose of benchmarking is to facilitate comparisons across firms. This paper provides industry practitioners with a range of benchmarks to evaluate the efficacy of models for operational loss projections within the Comprehensive Capital Analysis and Review (CCAR).

CCAR is an annual exercise by the Federal Reserve to assess whether the largest bank holding companies (BHCs) operating in the United States have sufficient capital to continue operations throughout times of economic and financial stress and that they have robust, forward-looking capital-planning processes that account for their unique risks. As part of the CCAR exercise, BHCs are required to project operational losses over a nine-quarter horizon assuming various economic scenarios. However, CCAR-participating BHCs have struggled to find meaningful relationships between operational losses and the macroeconomy (Board of Governors of the Federal Reserve System 2015).[5] The uncertainty associated with estimating links between the macroeconomy and operational risk is inherently high due to large and infrequent loss events dominating operational risk exposure. This uncertainty is compounded by the short length of the datasets used for estimation and by the difficulty in defining appropriate dates for operational loss events. Some types of operational loss event, such as internal fraud, may occur undetected over extended periods. While other types of operational loss event, such as legal cases, often result in payouts years after the occurrence date. Therefore, the operational risk models used in stress testing are often sensitive to modeling assumptions and data changes. To mitigate these challenges, firms typically use benchmarks to assess the reasonableness of their loss projections and reduce model risk.

Assessing the appropriateness of operational loss projections is even more challenging when projections rely on subjective assessments, such as scenario analysis. Federal Reserve stress testing guidance advises banks to focus on linking operational loss projections to risk identification processes and to use scenario analysis, while it discourages reliance on "unintuitive correlations" or on "distribution-based approaches that rely on historical data and require significant assumptions when projecting large operational losses" (Board of Governors of the Federal Reserve System 2015). In operational risk, scenarios analysis typically involves using risk assessments obtained from business and risk experts in workshops to project exposure. Banks adopt approaches aimed at making this process unbiased; nevertheless, estimates are largely driven by opinion and thus inherently subjective. Increased reliance on subjective approaches to project operational losses presents new challenges to operational risk managers. In particular, assessing the conservativeness of loss projections based on subjective or qualitative methods is more challenging than for quantitative models. Quantitative models typically enable the modeler to set the desired degree of

[5] Multiple researchers have explored these relationships in recent years. Allen and Bali (2007) found evidence of pro-cyclicality in operational losses using equity returns of financial institutions. Similarly, Cope et al. (2012) find a significant, positive relationship between GDP per capita and operational losses associated with Basel II event types in external fraud (EF) and employment practices and workplace safety (EPWS) in a large dataset from the Operational Riskdata eXchange (ORX) consortium. Lastly, Abdymomunov et al. (2017) analyzes US Federal Reserve supervisory data and finds that operational loss frequency and severity increases during economic downturns.

conservatism (e.g., the confidence level of the estimate) and can more easily be subject to empirical assessments of conservatism such as backtesting. Qualitative methods such as scenario analysis can aim for different levels of conservatism, but implementing such distinctions is challenging as they depend on experts ability to differentiate subjectively between different levels of conservatism (e.g., differentiating between a one-in-ten years event and a one-in-twenty years event). Also, implementation of backtesting or other outcome analysis for qualitative approaches is typically more challenging, as it requires a history of similarly produced qualitative model outputs.

In this paper, we try to fill this gap by providing practitioners a multiplicity of benchmarks that can be used to understand the conservatism of operational loss projections. We propose three types of benchmarks for the operational loss models of US BHCs: financial statement benchmarks, loss history benchmarks, and benchmarks that combine financial statement and loss history. We consider three financial statement benchmarks: total assets, risk-weighted assets (RWA), and gross income (GI). These benchmarks are simple and can be calculated with publicly available information. Also, previous research has documented the relation between scale of activity and operational loss exposure (Chernobai et al. 2011, Abdymomunov and Curti 2017, Curti and Migueis 2017). Thus, these size and business volume metrics are our starting point to benchmark banks' operational loss projections. In using these financial statement metrics to benchmark operational loss projections, we calculate for each BHC the ratios between operational loss projections and the financial statement benchmarks and provide descriptive statistics of these ratios. The median ratio of operational loss projections under the BHC stress scenario to total assets is approximately 0.7% of total assets and the median ratio to GI is approximately 15.5%. In addition, we calculate the historical distribution of nine-quarter operational losses to total assets and gross income, and display these distributions graphically. The 95th percentile of the historical nine-consecutive-quarter operational loss to total assets ratio is approximately 0.8% and the 95th percentile of the historical nine-consecutive-quarter operational loss to GI ratio is approximately 18.6%.

Following on research showing that historical operational loss experience is predictive of future operational loss exposure (Curti and Migueis 2017), we consider a variety of benchmarks based on historical loss history. First, we present benchmarks based on descriptive statistics (average, median, and maximum) of historical nine-consecutive-quarter total losses and historical nine-consecutive-quarter loss frequency. The median ratio of BHC stress operational loss projections to average nine-quarter total operational losses is 4.2, to median nine-quarter total operational losses is 6.3, and to maximum nine-quarter total operational losses is 1.5. Second, we present benchmarks based on a simplified loss distribution approach (LDA) model, the empirical bootstrap, for which we calculate multiple percentiles of the nine-quarter operational loss distribution. The median ratio of BHC stress operational loss projections to the 95th percentile of the nine-quarter loss distribution obtained through empirical bootstrap is approximately two.

Finally, we present benchmarks based on both a financial statement metric (total assets) and a historical loss metric (average nine-consecutive-quarter loss frequency). Curti and Migueis (2017) showed that loss frequency is the historical loss metric that best predicts tail operational loss exposure and that prediction accuracy is maximized when loss frequency is combined with controls for firm size. Thus, following that paper, we generate benchmarks for multiple percentiles by using average loss frequency together with asset size within a quantile regression framework. The median ratio of BHC stress losses to the 95th percentile estimate of the operational loss distribution obtained through quantile regression is 0.9.

Correlation analysis of benchmark ratios shows that while some pairs of benchmarks are highly correlated, other pairs have little correlation. The proposed benchmarks provide different proxies of operational loss exposure and thus are helpful to provide a comprehensive view of the conservativeness of projected losses.

The remainder of this paper is organized as follows: Section II describes the data used in the analysis; Section III presents the financial statement benchmarks; Section IV presents the loss history benchmarks; Section V presents the benchmarks that combine financial statement and loss information; Section VI discusses the correlation between the alternative benchmarks; finally, Section VII concludes.

II – Data

As of 2018, BHCs with assets over $50 billion were required to provide operational loss data and operational loss projections, under a variety of scenarios, to the Federal Reserve System as part of CCAR. This data is used in our analysis. At the time of analysis, data was submitted by 38 BHCs. CCAR BHCs must report information on all their operational loss events above an appropriate collection threshold including dollar amount and accounting date of the loss. Also, CCAR BHCs must report operational loss projections under a "BHC stressed scenario."[6] This scenario is meant to focus on BHCs specific vulnerabilities and, thus, is generally the most severe scenario for operational losses. For this reason, we focus on it in this paper. Finally, the analysis uses financial statement variables collected from the FR Y-9C reports, including total assets, RWA, and GI.

The operational loss data available varies across BHCs. BHCs' participation in CCAR started in different points in time, and so certain BHCs started their loss collection fairly recently. Other BHCs, generally those subject to the advanced measurement approaches (AMA) for operational risk capital, have collected operational loss data since 2000 (or even earlier). In our analysis, we include loss data starting from 2000Q1 or, if data is only available more recently, as far back as data is reported in the FR-14Q. The loss collection thresholds also differ across BHCs. To achieve comparability in our historical loss benchmarks, we only include operational loss events above $20,000 in loss totals and on the estimation of the bootstrapped operational loss distribution. As of 2018, the loss collection threshold is $20,000 or lower for all CCAR BHCs. Gross loss amounts are used, adjusted for inflation.[7] All impacts of an operational loss event are aggregate into the first quarter in which the event produced an accounting impact.

[6] For a detailed description of the operational loss data and the BHC stress loss projections provided by BHCs to the Federal Reserve, see the FR Y-14Q and FR Y-14A reporting forms and instructions at www.federalreserve.gov/apps/reportforms.

[7] Losses are adjusted using the Personal Consumption Expenditures Excluding Food and Energy Series obtained from the Federal Reserve Economic Data (FRED).

III – Financial Statement Benchmarks

We consider three alternative financial statement metrics to benchmark stress loss projections: total assets, RWA, and GI.

Abdymomunov and Curti (2017) showed that operational losses are linked to the asset size of banks. Therefore, we believe benchmarking operational loss projections to total assets is useful. Nevertheless, this benchmark has limitations. BHCs have different degrees of off-balance sheet exposure (e.g., derivatives), which are not represented in total assets despite these activities often resulting in large operational risk exposure. Also, certain businesses with high operational risk (e.g., underwriting, securitization) may not result in large asset holdings. For these reasons, practitioners should exercise caution when using total assets to compare operational loss projections across firms.

The total assets metric does not account for differences in asset riskiness. To address this limitation, the Basel accords created RWA, where assets of different riskiness are weighted differently. To understand how operational loss projections compare to the overall risk profile of the firm, we benchmark operational loss projections to RWA. For comparability across firms, we use standardized approach RWA, which are available for all firms, and do not use advanced approaches RWA for advanced approaches firms.

Finally, we use GI to benchmark operational loss projections. GI is the proxy for operational risk used in Basel II's standardized approaches. To be consistent with the Basel Committee's use of the GI, we use average GI over a three-year window to benchmark operational loss projections. So, first we calculate GI for each year by summing net interest income and total noninterest income from the schedule HI of the FR Y-9C report; then, we average GI figures for a rolling window of three years (e.g., to obtain average 2017 GI, we average GI for 2015, 2016, and 2017). Similar to asset size, risk managers should exercise caution in using GI-scaling to compare their loss projections with peers because GI captures profitability, which may not always correlate with the volume of activities generating operational risk. For example, a high trading volume firm, with low GI due to trading losses, may still have large operational loss exposure.

These three benchmarks provide alternative views to compare operational loss projections across banks and through time. All of them are simple and can be obtained from public reports. Table 1 displays descriptive statistics for the ratio of BHC stressed loss projections to the three financial statement metrics, as well as the correlation between BHC stress loss projections and the three metrics.

Table 1

	Ratio of BHC Stress Losses to		
Descriptive Statistics	**Total Assets**	**RWA**	**GI**
10th Percentile	0.3%	0.4%	7.0%
25th Percentile	0.5%	0.7%	12.1%
Median	0.7%	1.0%	15.5%
75th Percentile	1.0%	1.6%	21.4%
90th Percentile	1.5%	3.5%	32.6%
Average	0.8%	1.4%	18.7%
	Correlation of BHC Stress Losses with		
	Total Assets	**RWA**	**GI**
	97.0%	96.1%	95.0%
Notes: N = 38.			

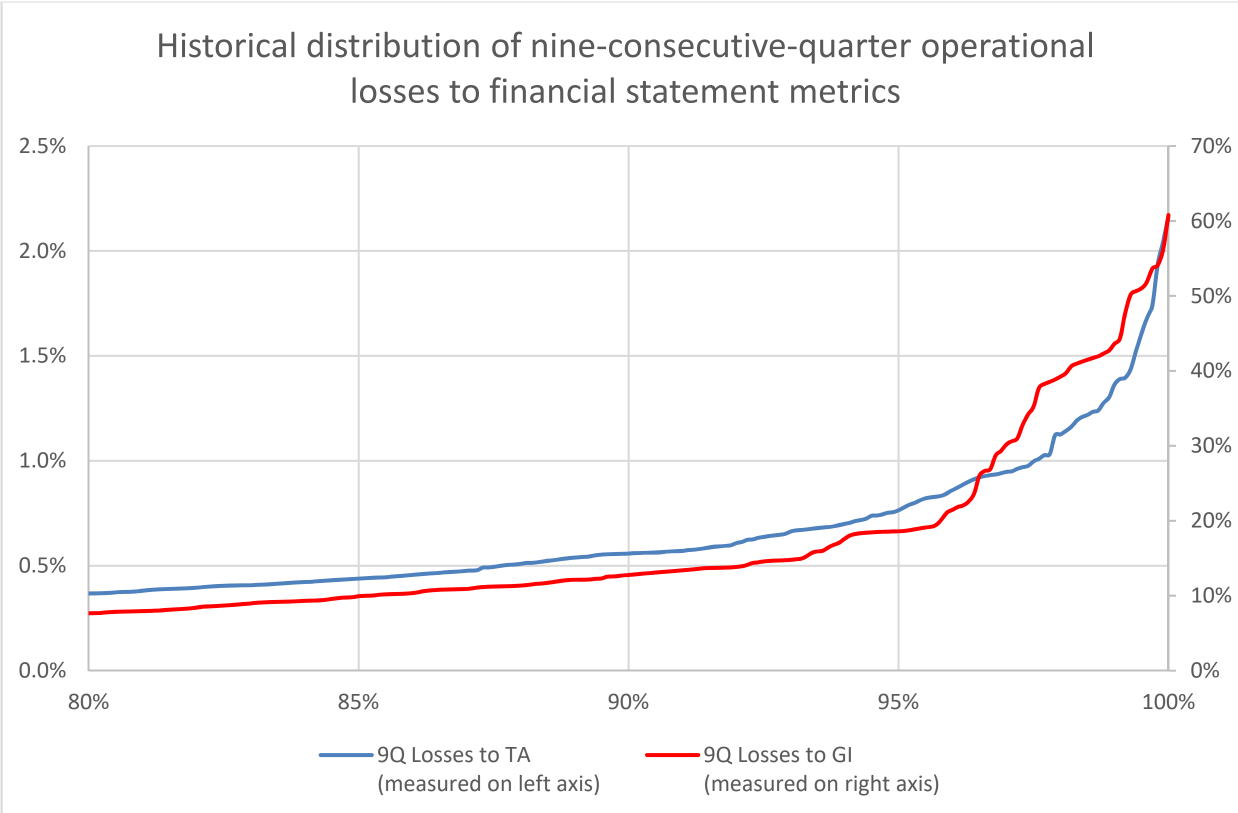
The ratio of BHC stressed operational losses to total assets is wide-ranging, with the 10th percentile at 0.3%, the industry median at 0.7%, and the 90th percentile at 1.5%. While the ratio of BHC stressed operational losses to risk-weighted assets ranges from 0.3% at the 10th percentile to 4.9% at the 90th percentile, with an industry median of 1%. The Basel Accords require banks to hold capital to, at the minimum, cover losses amounting to 8% of RWA. The ratio between stressed operational losses and these minimum capital requirements show how significant operational risk is. The median bank projects that, over a nine quarter stress environment, operational losses may erode approximately 12% of total risk-based capital;[8] while for 25% of banks operational losses may erode more than 20% of total risk-based capital, and for 10% of banks more than 44% of risk-based capital. Finally, stress loss projections range from 7% of GI at the 10th percentile to 32.6% of GI at the 90th percentile, with an industry median of 15.5%. Under Basel II, banks could use the Basic Indicator Approach (BIA) to set operational risk capital at 15% of GI. For more than half of CCAR BHCs, projected stress operational losses exceed 15% of GI.

Of total assets, RWA, and GI, BHC stressed operational losses are most correlated with total assets (97%). We believe all three benchmarks are useful, as they allow for comparisons across different dimensions. Nevertheless, the higher correlation of total assets with projected losses indicates that total assets best track firm projections of operational loss exposure.

To assess the degree of conservatism of stress loss projections, we calculate the historical distribution of the ratio of nine-consecutive-quarter operational losses to total assets and of the ratio of nine-consecutive-quarter operational losses to GI. By taking into account the overall industry experience with operational losses, the distributions of these ratios illustrate the potential size of tail losses for banks of a given asset size or GI. Figure 1 displays the values of the ratios above the 80th percentile of nine-consecutive-quarter losses to total assets and of nine-consecutive-quarter losses to GI.

[8] Basel required total capital is equal to 1/0.08 times RWA. Thus, the loss projections of the BHC with the median projections to RWA ratio (0.994%) translate into approximately a 12% reduction in RWA [≈ 0.994%*(1/0.08)].

Figure 1



At the 95th percentile, nine-consecutive-quarter operational losses reach 0.76% of total assets and 18.6% of GI. Losses increase at an accelerating rate as we move toward the extreme tail of the distribution, reaching 1.36% of total assets and 43.6% of GI at the 99th percentile. The median BHC's stress loss projections are at the 93.9th percentile of the industry distribution of the nine-consecutive-quarter loss to total assets ratio and at the 93.4th percentile when the ratio to GI is used.

IV – Loss History Benchmarks

To understand and compare banks' operational risk loss exposure, loss history is considered next. Past loss experience is a reasonable proxy for operational risk loss exposure going forward because past losses relate to the quality of risk controls and the riskiness of the business environment (Curti and Migueis 2017). However, caution should be applied when making comparisons across firms based on internal loss benchmarks because of their disparate data collection history and quality. Some BHCs have collected more than eighteen years of data and have a mature collection process, while others have less than ten years of data. Generally, the BHCs subject to Supervision and Regulation (SR) Letter 15-18 have more complete operational risk datasets.

We consider three types of benchmarks based on past losses: 1) statistics of historical nine-consecutive-quarter total losses; 2) statistics of historical nine-consecutive-quarter loss frequency; and 3) tail percentiles of the nine-quarter operational loss distribution, calculated through empirical bootstrapping.

Nine-consecutive-quarter total losses

Using ratios between BHC stress loss projections to descriptive statistics of historical loss totals is one the simplest ways to benchmark projections across firms. We calculate three statistics of the historical distribution of nine-consecutive-quarter total losses for each BHC: median, mean, and maximum. Then, for each BHC, we calculate ratios of BHC stress loss projections to each of these three statistics. Finally, we calculate the descriptive statistics of these ratios for the 38 BHCs in our sample. Table 2 presents these industry descriptive statistics plus the correlation of BHC stress loss projections with the mean, median, and maximum of nine-consecutive-quarter total losses.

Table 2

	Ratio of BHC Stress Losses to		
Descriptive Statistics	**Median 9Q Total Loss**	**Mean 9Q Total Loss**	**Maximum 9Q Total Loss**
10th Percentile	2.4	1.8	0.5
25th Percentile	3.6	2.6	1.0
Median	6.3	4.2	1.5
75th Percentile	8.8	7.3	2.8
90th Percentile	16.4	10.4	6.3
Average	7.4	5.8	2.7
	Correlation of BHC Stress Losses with		
	Median 9Q Total Loss	**Mean 9Q Total Loss**	**Maximum 9Q Total Loss**
	90.0%	91.8%	87.8%
Notes: N = 38.			

All BHCs have stress loss projections above their median and mean nine-consecutive-quarter total losses. Nevertheless, the dispersion of these ratios across firms is large. Ten percent of firms have a ratio of stress loss projections to mean losses below 1.8, while ten percent of firms have a ratio above 10.4. The ratio of BHC stress loss projections to maximum nine-consecutive-quarter total losses is 1.5 for the median firm, indicating that the median firm projects operational losses under such stress scenario to be 50% larger than the worst historical experience. Meanwhile, BHC stress loss projections are smaller than maximum nine-consecutive-quarter total losses for 25% of firms. This indicates that a meaningful proportion of firms project that stress losses would fall below their worst historical experience.

Correlation analysis indicates that BHC stress loss projections correlate the most with mean losses and the least with maximum losses. The smaller correlation projections with maximum losses is understandable as past maximum losses may have resulted from idiosyncratic events that are unlikely to be repeated in the future. Note that all three statistics of nine-consecutive-quarter total losses correlate less with BHC stress loss projections than the measures of size (total assets, RWA, and GI) considered in the previous section. This seems to indicate that firms consider their size more than their historical loss experience when projecting stress loss exposure.

Nine-consecutive-quarter loss frequency

Curti and Migueis (2017) showed that loss frequency is predictive of operational loss exposure. Moreover, that paper finds that average loss frequency is the historical loss metric most predictive of future losses out of multiple alternatives. The better performance of loss frequency is likely due to it being a more stable proxy for firms' control failures than other loss history metrics, such as average total losses. For these reasons, we also consider benchmarks based on banks' historical loss frequency.

Similar to loss totals, we benchmark BHCs relative to three descriptive statistics of nine-consecutive-quarter loss frequency: median, mean, and maximum. We only include loss events above $20,000 because this is the smallest amount for which all BHCs have collected data. Table 3 presents industry descriptive statistics for the ratio of BHC stress loss projections to each of these three statistics of loss frequency, as well as the correlation between BHC stress loss projections and each of these three statistics of loss frequency.

Table 3

	Ratio of BHC Stress Losses ($ millions) to		
Descriptive Statistics	**Median 9Q Loss Frequency**	**Mean 9Q Loss Frequency**	**Maximum 9Q Loss Frequency**
10th Percentile	1.2	1.1	0.7
25th Percentile	1.5	1.5	1.1
Median	2.9	2.6	1.6
75th Percentile	4.5	4.3	2.7
90th Percentile	8.1	7.4	4.5
Average	3.8	3.7	2.3
	Correlation of BHC Stress Losses with		
	Median 9Q Loss Frequency	**Mean 9Q Loss Frequency**	**Maximum 9Q Loss Frequency**
	94.3%	95.2%	95.6%
Notes: N = 38.			

The correlation of BHC stress loss projections to median, mean, and maximum nine-consecutive-quarter loss frequency is meaningfully higher than the correlation observed to median, mean, and minimum nine-consecutive-quarter total losses. So, just like Curti and Migueis (2017) found that historical loss frequency is a better predictor of future exposure than historical loss totals, BHCs' own stress loss projections tie closer to their historical loss frequency than to their historical loss totals.

Nevertheless, the ratio of BHC stress loss projections to loss frequency statistics displays meaningful dispersion across firms. For example, the ratio of BHC stress loss projections to historical nine-consecutive-quarter mean loss frequency ranges from $1.1 million at the 10th percentile to $7.4 million at the 90th percentile.

Tail quantiles of the nine-quarter operational loss distribution (empirical bootstrapping)

An alternative approach to benchmark BHC stress loss projections is to follow a simplified LDA framework to estimate tail percentiles for the nine-quarter operational loss distribution of each BHC, using each BHC's operational loss history. The Federal Reserve has recommended this simplified LDA framework, commonly known as the "empirical bootstrap," as a benchmark to AMA models in recent supervisory guidance (Board of Governors of the Federal Reserve System 2014).

Our implementation of the empirical bootstrap begins with separating each BHC's losses into the seven Basel event types, which are used as units of measure. Then, loss frequency and severity are modeled separately and assumed independent. We assume loss frequency follows a Poisson distribution. Regarding severity, we follow a non-parametric approach by setting the severity distribution equal to the historical empirical distribution of observed loss severities.

After obtaining the frequency and severity distribution of each event type, a Monte Carlo simulation is used to obtain the distribution of aggregate total losses over a nine-quarter period. Each simulation path proceeds as follows: first, a number of loss events is drawn from the Poisson distribution for each event type; second, the corresponding number of loss severities are drawn from the empirical severity distribution of each even type; third, these severities are added up to obtain the aggregate loss for the event type in a simulated nine-quarter period; fourth, we sum the aggregate loss of the seven event types to get one simulated nine-quarter total operational loss for the BHC (i.e., we assume independency across event types). We repeat this procedure 100,000 times to obtain the distribution of nine-quarter total operational losses for a BHC.

The empirical bootstrap we adopted likely underestimates exposure at high quantiles for three reasons: 1) the empirical bootstrap assumes that the severity of a loss event can never exceed the largest loss historically observed and, thus, given the relatively short operational loss datasets of most banks, it likely underestimates the tail percentiles of the underlying operational loss distribution; 2) there is likely positive dependency across event types (Cope and Antonini 2008, Abdymomunov and Ergen 2017) which we do not capture; and 3) both loss severity and loss frequency increase with economic distress (Abdymomunov et al. 2017) and thus are positively correlated in the tail region of their distributions. Nevertheless, we did not try to address these shortcomings because correcting them would require adding complications to the estimation. The main purpose of this benchmark is to facilitate comparison of projections across the industry rather than to obtain the best possible estimate of each BHC's tail exposure, and this objective is best achieved if the benchmark is kept as simple as possible.

We consider three alternative tail percentiles as benchmarks: the 90th, the 95th, and the 99th. Table 4 provides descriptive statistics for the ratio of BHC stress loss projections to these percentiles of the empirical bootstrap, as well as the correlation between BHC stress loss projections and the percentiles of the empirical bootstrap.

Table 4

	Ratio of BHC Stress Losses to		
Descriptive Statistics	**90th Percentile Emp Bootstrap**	**95th Percentile Emp Bootstrap**	**99th Percentile Emp Bootstrap**
10th Percentile	0.8	0.7	0.5
25th Percentile	1.4	1.2	0.8
Median	2.2	2.0	1.4
75th Percentile	3.9	3.4	2.4
90th Percentile	6.9	6.3	5.3
Average	3.5	3.1	2.3
	Correlation of BHC Stress Losses with		
	90th Percentile Emp Bootstrap	**95th Percentile Emp Bootstrap**	**99th Percentile Emp Bootstrap**
	90.5%	90.3%	90.0%
Notes: N = 38.			

BHC stress loss projections cover the 90th and the 95th percentiles of the empirical bootstrap nine-quarter operational loss distribution for 84% of the BHCs and the 99th percentile for 63% of the BHCs. Meanwhile, the ratio of BHC stress loss projections to the 99th percentile of the empirical bootstrap operational loss distribution is 1.4 for the median firm. These statistics imply that, according to the empirical bootstrap, most firms are projecting their BHC stress losses to be higher than a 99th percentile, a one in a hundred years event. However, as previously discussed, the empirical bootstrap benchmark we have implemented likely underestimates the tail region of the operational loss distribution.

As we showed in the section discussing benchmarks based on the financial statement, a comparison between the BHC stress projections and the industry-wide historical distribution of the ratio between losses and size metrics shows that median projections are somewhat below the 95th percentile of this distribution. Besides the likely underestimation of exposure under the bootstrapping approach, the different assessment of the conservatism of projections resulting from the two approaches is likely due to the largest operational losses (in proportion to size) having been concentrated in a few firms. Such losses influence the industry wide-distribution of the loss-to-size ratio, but are not reflected in the empirical bootstrap of the firms that did not experience them. Judging which of the approaches better reflects the degree of conservatism of BHC projections is difficult, as the answer depends on whether the potential for large losses is truly concentrated in the institutions that experienced them in the past (in which case the empirical bootstrap is a preferable approach to assess conservatism) or whether, under different circumstances, large losses could have occurred to other firms of similar size (in which case the size-based benchmarks are better to assess conservatism).

The correlation between BHC stress loss projections and tail percentiles obtained through the empirical bootstrap is approximately 90%. This correlation is similar to the values obtained for average and median total losses, but falls short of the correlation of BHC stress loss projections to measures of size or of loss frequency.

V – Financial Statement and Loss History Benchmarks

As shown so far, BHC stress loss projections correlate closely both with metrics of firm size/business volume and with metrics of loss history. Among these, asset size and loss frequency are most correlated with BHC stress loss projections. In addition, Curti and Migueis (2017) showed that tail exposure is best predicted when loss history and size metrics are combined. For these reasons, we also produce benchmarks for BHC stress loss projections based on the quantile regressions where total assets and average loss frequency are used together as explanatory variables. This approach yields benchmarks that are conceptually similar to the new standardized approach for operational risk set by the Basel Committee (Basel Committee on Banking Supervision 2017), albeit calibrated to the tail experience of US BHCs.

We follow the quantile regression methodology outlined in Curti and Migueis (2017), but do not include other explanatory variables besides total assets and average nine-consecutive-quarter loss frequency. Similar to the empirical bootstrap analysis, we consider three alternative percentiles for the quantile regressions: 90^{th}, 95^{th}, and 99^{th}. The equations below present the results of these three quantile regressions:

$$90^{th} percentile\ 9Q\ operational\ loss\ distribution \\ = 0.52\% \cdot Total\ Assets + 435{,}376 \cdot Avg\ 9Q\ Loss\ Frequency$$

$$95^{th} percentile\ 9Q\ operational\ loss\ distribution \\ = 0.80\% \cdot Total\ Assets + 237{,}688 \cdot Avg\ 9Q\ Loss\ Frequency$$

$$99^{th} percentile\ 9Q\ operational\ loss\ distribution \\ = 1.00\% \cdot Total\ Assets + 2{,}213{,}062 \cdot Avg\ 9Q\ Loss\ Frequency$$

Table 5 provides descriptive statistics for the ratio of BHC stress loss projections to the three quantile projections, as well as the correlation between BHC stress loss projections and the three quantile projections.

Table 5

	Ratio of BHC Stress Losses to		
Descriptive Statistics	**90^{th} Percentile Quantile Reg**	**95^{th} Percentile Quantile Reg**	**99^{th} Percentile Quantile Reg**
10^{th} Percentile	0.5	0.4	0.2
25^{th} Percentile	0.8	0.6	0.3
Median	1.1	0.8	0.4
75^{th} Percentile	1.4	1.1	0.5
90^{th} Percentile	2.2	1.7	0.9
Average	1.3	0.9	0.5
	Correlation of BHC Stress Losses with		
	90^{th} Percentile Quantile Reg	**95^{th} Percentile Quantile Reg**	**99^{th} Percentile Quantile Reg**
	97.8%	97.4%	97.9%
Notes: N = 38.			

BHC stress loss projections cover the 90^{th} percentile of the nine-quarter operational loss distribution obtained through quantile regression for 58% of the BHCs, the 95^{th} percentile for 29% of the BHCs, and the 99^{th} percentile for 8% of the BHCs. For the median firm, BHC stress loss projections correspond to a percentile between the 90^{th} and 95^{th} under this quantile regression methodology. The degree of conservatism of BHC stress projections implied by this approach is similar to the degree of conservatism implied by benchmarking to financial statement metrics alone. This happens because, similar to the historical ratio exercise we undertook to assess the conservatism of projections using assets and GI, the quantile regression approach considers the full industry sample in projecting tail exposure. Thus, large losses relative to size observed for some firms are used within this approach to inform the tail losses others firms may experience in the future.

The correlation of BHC stress loss projections to these tail percentiles obtained from quantile regression is the highest of any of the benchmarks considered in this paper. This finding suggests that BHCs projections reflect a mix of their size and their historical loss experience.

VI – Correlation across Benchmarks

To assess whether the benchmarks in this paper provide alternative views into the conservatism of BHCs' projections, we calculated the pairwise correlation between benchmark ratios for all pairs of benchmarks. Some pairs of benchmark ratios are highly correlated, but other pairs are not. Table 6 presents these correlations. Generally, there are two groups of benchmarks where all benchmarks are highly correlated with each other, but have little correlation with benchmarks in the other group: 1) ratios based on asset size, income, quantile regression percentiles, and loss frequency; and 2) ratios based on total loss statistics and empirical bootstrap percentiles.

The high correlation among ratios based on size metrics is unsurprising. Also, their high correlation with the ratios based on quantile regression percentiles is explained by total assets being the main driver of the quantile regression estimates. Meanwhile, the high correlation between ratios based on loss total statistics and ratios based on empirical bootstrap percentiles is likely due to empirical bootstrap estimates being driven by the large losses in a BHCs loss history, just like the maximum nine-consecutive-quarter loss and, to a large extent, the mean nine-consecutive-quarter loss are.

The higher correlation of ratios based on loss frequency statistics with size-based ratios than with ratios based on loss total statistics may be surprising. But it illustrates that while loss averages and LDA-like models, such as the empirical bootstrap, are driven by large losses (which often do not correlate strongly with size), loss frequency tracks much closer to firm size.

This correlation analysis shows there are two main sources of variation captured by the benchmarks we propose: historical experience of large losses and firm size. In our view, given the uncertainty over how much of operational risk exposure is idiosyncratic to a firm vs. systemic to the industry, considering benchmarks that represent both views is relevant. Also, while some of the pairwise benchmark correlations within the two main groups are high, the ranking they produce still has meaningful differences for some firms and thus we believe validators can gain useful insights from examining the various benchmarks proposed in this paper when assessing the conservatism of their firms' projections.

VII – Conclusion

This paper proposes multiple benchmarks for the stress operational loss projections of large US BHCs. Following previous research documenting the relation of operational loss exposure to firm size, we benchmark BHC loss projections to total assets, RWA, and GI. Also, following research showing that operational loss exposure is persistent and, thus, that past operational losses are predictive of future losses, we use summary statistics of historical total losses and historical loss frequency to benchmark stress loss projections. Finally, we build two model-based benchmarks, based on the empirical bootstrapping approach and on quantile regression, to assess the conservatism of BHC stress loss projections.

Using three different approaches, we find a range for the theoretical conservatism of BHC stress loss projections. The two approaches that pool industry information to assess the conservatism of approaches (our historical loss-to-size ratio approach and the quantile regression) find that the median BHC's stress loss projections fall somewhere between the 90th and 95th percentile of their operational loss distribution. On the other hand, the empirical bootstrap approach, which focuses on a BHC's own loss distribution, finds that the median BHC's stress loss projections are above the 99th percentile of their operational loss distribution.

Correlation analysis shows that both size metrics and historical loss metrics are highly correlated with BHC stress loss projections. But correlation is generally higher for size metrics. Among the loss metrics, historical loss frequency is the most correlated with loss projections. Meanwhile, the quantile regression benchmarks, which combine size and loss frequency, have the highest correlation with BHC's stress loss projections. These correlations suggest that firms consider both size and past loss experience when projecting stress operational losses.

We also performed correlation analysis across the benchmark ratios and found that while some of the benchmark-ratio pairs are highly correlated, other pairs are not, which indicates that the array of benchmarks proposed in this paper provide different views into operational loss exposure. Firms can use the various benchmarks provided in this paper to assess whether the conservatism of their projections is reasonable given their risk appetite. In particular, firms whose stress loss projections are in the bottom of the distribution across the various benchmarks should consider whether their projections are appropriately conservative.

We believe the set of benchmarks presented in this paper can form the basis for a robust benchmarking framework. Nevertheless, benchmarks should be re-evaluated as models, regulations, and economic factors change. A sound operational risk modeling framework combined with ample use of benchmarks offers the best way forward for measuring and managing operational risk.

Table 6

	Correlation between ratio of BHC stress loss projections to row variable and ratio of BHC stress loss projections to column variable														
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Assets (1)	1.00														
RWA (2)	0.89	1.00													
GI (3)	0.81	0.76	1.00												
Total Loss Median (4)	0.07	-0.09	0.16	1.00											
Total Loss Mean (5)	-0.04	-0.14	0.06	0.88	1.00										
Total Loss Max (6)	-0.05	-0.15	0.05	0.77	0.96	1.00									
Frequency Median (7)	0.68	0.59	0.71	0.38	0.20	0.14	1.00								
Frequency Mean (8)	0.71	0.62	0.73	0.36	0.22	0.17	0.98	1.00							
Frequency Max (9)	0.66	0.56	0.67	0.40	0.32	0.29	0.92	0.97	1.00						
Bootstrap 90^{th} (10)	-0.08	-0.16	0.03	0.77	0.97	0.96	0.09	0.12	0.23	1.00					
Bootstrap 95^{th} (11)	-0.08	-0.16	0.03	0.77	0.97	0.96	0.09	0.12	0.23	1.00	1.00				
Bootstrap 99^{th} (12)	-0.08	-0.16	0.02	0.73	0.95	0.95	0.07	0.10	0.20	1.00	1.00	1.00			
Quantile Reg 90^{th} (13)	0.97	0.87	0.84	0.17	0.05	0.03	0.80	0.83	0.79	-0.02	-0.01	-0.03	1.00		
Quantile Reg 95^{th} (14)	0.99	0.89	0.83	0.11	0.00	-0.01	0.74	0.77	0.72	-0.05	-0.05	-0.06	0.99	1.00	
Quantile Reg 99^{th} (15)	0.92	0.83	0.83	0.24	0.11	0.08	0.87	0.90	0.87	0.03	0.03	0.01	0.99	0.96	1.00

Notes: The numbers in parenthesis in the column headers correspond to the variable numbers indicated in row headers. N = 38.

References

Abdymomunov, Azamat, and Filippo Curti (2017). Improving the Robustness of Operational Risk Estimation and Stress Testing Through External Data Scaling. Social Science Research Network.

Abdymomunov, Azamat, Filippo Curti, and Atanas Mihov (2017). U.S. Banking Sector Operational Losses and the Macroeconomic Environment. Social Science Research Network.

Abdymomunov, Azamat, and Ibrahim Ergen (2017). Tail Dependence and Systemic Risk in Operational Losses of the US Banking Industry. *International Review of Finance* Vol 17, Issue 2: 177-204.

Allen, Linda, and Turan Bali (2007). Cyclicality in catastrophic and operational risk measurements. *Journal of Banking and Finance* 31: 1191–1235.

Basel Committee on Banking Supervision (2017). Basel III: Finalising post-crisis reforms.

Board of Governors of the Federal Reserve System (2011). SR 11-7: Guidance on Model Risk Management.

Board of Governors of the Federal Reserve System (2014). BCC 14-1: Supervisory Guidance for Data, Modeling, and Model Risk Management Under the Operational Risk Advanced Measurement Approaches.

Board of Governors of the Federal Reserve System (2015). SR 15-18: Federal Reserve Supervisory Assessment of Capital Planning and Positions for LISCC Firms and Large and Complex Firms.

Chernobai, Anna, Philippe Jorion, and Fan Yu (2011). The Determinants of Operational Risk in U.S. Financial Institutions. *Journal of Financial and Quantitative Analysis* Vol. 46, No. 6: 1683–1725.

Cope, Eric, and Gianluca Antonini (2008). Observed correlations and dependencies among operational losses in the ORX consortium database. *The Journal of Operational Risk* Vol. 3, No. 4: 47-74.

Cope, Eric, Mark Piche, and John Walter (2012). Macroenvironmental determinants of operational loss severity. *Journal of Banking and Finance* 36: 1362-1380.

Curti, Filippo, and Marco Migueis (2017). Predicting Operational Loss Exposure Using Past Losses. Social Science Research Network.